[19]中华人民共和国专利局

[11] 公开号 CN 1138401A

[12]发明专利申请公开说明书

[21]申请号 95191135.X

[51]Int.Cl^6

H04N 1/415

[43]公开日 1996年12月18日

[22]申请日 95.8.31

[30]优先权

[32]94.11.4 [33]US[31]08/334,718

[86]国际申请 PCT/US95/11187 95.8.31

[87]国际公布 WO96/14708 英 96.5.17

[85]进入国家阶段日期 96.7.2

[71]申请人 摩托罗拉公司

地址 美国伊利诺斯

[72]发明人 詹姆斯·查尔斯·布雷恩

[74]专利代理机构 中国国际贸易促进委员会专利商标事务所

代理人 杨晓光

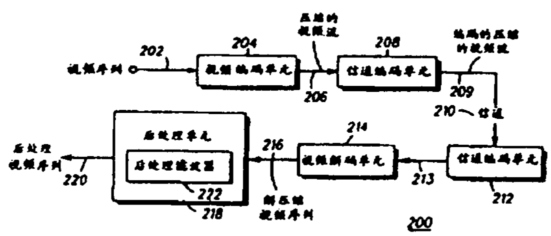
G06T 5/00

权利要求书 5 页 说明书 10 页 附图页数 3 页

[54]发明名称 用于抑制人工形成物的方法,后处理滤波器与视频压缩系统

[57]摘要

通过计算后处理象素强度有效抑制视频压缩系统(200)中的蚊形和分块人工形成物的可见性。应用预定的噪声方差、预定的自相关系数、和包含多个先前后处理象素强度和多个原有的象素强度的象素的一个局部邻域确定(100)后处理象素强度。这消除了对于局部信号和噪声功率的估计值的依赖性。



权 利 要 求 书

1. 一种使用滤波器来获取后处理的象素强度的方法，用于驻留在存储器中的解压缩的视频序列中的帧，该滤波器基于预定的噪声方差 和 预定的自相关系数而抑制蚊形及分块人工形成物的可见性，该方法包括步骤：

1A）在一个存储器单元中，基于解压缩的视频序列而生成象素的一个局部邻域，其中该象素的局部邻域包括多个先前的后处理的象素强度以及多个解压缩的象素强度；

1B）基于象素的局部邻域估算信号方差；

1C）基于信号方差、预定的噪声方差、及预定的自相关系数计算一个优化的滤波器增益；以及

1D）应用一个加法器，基于预定的自相关系数、象素的局部邻域、以及优化的滤波器增益确定后处理象素强度。

2. 权利要求 1 的方法，其中至少 2A—2C 之一成立：

2A）预定的噪声方差为 1 到 10 之间的一个数值；

2B）应用形式为：

$$\hat{\sigma}_s^2(n)=\frac{1}{2M+1}\left[\sum_{i=-M}^{0}(x(n-i))^2+\sum_{i=1}^{M}(\hat{y}(n-i))^2\right]-\hat{m}_s^2(n)$$

的方程式估算信号方差，其中

$$\hat{m}_s(n)=\frac{1}{2M+1}\left[\sum_{i=-M}^{0}x(n-i)+\sum_{i=1}^{M}\hat{y}(n-i)\right],$$

为象素的局部邻域中的平均强度值，$(2M+1)$是邻域尺寸，$\hat{y}(n)$是象素的局部邻域中先前后处理的象素强度；以及

2C）应用形式为：

$$G(n)=\frac{\alpha\bullet\sigma_s^2(n)}{\sigma_n^2+\alpha\bullet\sigma_s^2(n)}$$

的方程式计算优化的滤波器增益，其中

$$\alpha=1-\frac{2\bullet\zeta^2}{1+\zeta^2}$$

以及 σ_n^2 是预定的噪声方差，$\sigma_s^2(n)$是信号方差，而 ζ 是预定的自相关系数。

3. 权利要求 1 的方法，其中后处理象素强度进一步基于所选的象素的强度的估算值，并且其中选择 3A—3B 至少一项：

3A）其中所选象素的强度的估计值是预定的自相关系数、相对于预定的扫描方向下一个象素的解压缩象素的强度、以及紧靠前面的一个后处理象素强度的函数，并且此处进一步选择，其中所选象素的强度估计值应用形式为：

$$\hat{y}_a(n)=\frac{\zeta}{1+\zeta^2}\bullet\left[\hat{y}_b(n-1)+x(n+1)\right]$$

的方程式确定，其中 $\hat{y}_a(n)$为所选象素的强度的估计值，ζ 是预定的自相关系数，$x(n+1)$是相对于预定的扫描方向下一个象素的

解压缩象素的强度，以及 $\hat{y}_b(n-1)$ 是紧靠前面的一个后处理象素强度；以及

3B）后处理象素强度应用形式为：

$$\hat{y}_b(n)=\hat{y}_a(n)+G(n)\bullet\left[x(n)-\hat{y}_a(n)\right]$$

的方程式确定，其中 $G(n)$ 为优化滤波器增益，$\hat{y}_a(n)$ 是一象素的强度的估计值，而 $x(n)$ 是解压缩象素的强度。

4. 权利要求 1 的方法，其中后处理象素强度是基于多个预定的内插系数而被尺寸再设定的后处理象素强度，并且此处进一步选择，其中存储器为了确定尺寸再设定的后处理象素强度而被访问一次。

5. 用于抑制蚊形和分块人工形成物的可见性的后处理滤波器，包括：

5A）用于存储预定的自相关系数、预定的噪声方差、以及象素的局部邻域的一个存储器单元；

5B）能够与存储器单元耦合操作的一个信号方差确定器，用于基于象素的局部邻域估算信号方差；

5C）能够与存储器和信号方差确定器耦合操作的一个优化滤波器增益确定器，用于基于信号方差、预定的噪声方差、以及预定的自相关系数确定优化的滤波器增益；以及

5D）能够与存储器单元和优化滤波器增益确定器耦合操作的一个组合电路，用于基于预定的自相关系数、象素的局部邻域、以及优化的滤波器增益来确定后处理象素的强度。

6. 权利要求 5 的后处理滤波器，其中 6*A*—6*B* 至少一项成立：

6*A*）存储在存储器中的预定的噪声方差为 1 到 10 之间的一个数值；以及

6*B*）组合电路为有限脉冲响应滤波器。

7. 权利要求 5 的后处理滤波器，其中该组合电路还包括：

7*A*）能够与存储器单元耦合操作的一个预测器，用于预测当前象素的一个强度；以及

7*B*）能够与存储器单元预测器及优化的滤波器增益确定器耦合操作的一个更新器，用于确定后处理象素强度；

并且此处还选择，

7*C*）该预测器应用象素的局部邻域的一维非表因的自回归模型。

8. 权利要求 5 的后处理滤波器，其中组合电路是一个无限的脉冲响应滤波器。

9. 权利要求 5 的后处理滤波器，其中存储器单元还存储预定的多个内插系数，并且此处被选择，其中组合电路为一个尺寸再设定滤波器，并且此处还被选择，至少 9*A*—9*B* 之一：

9*A*）该组合电路访问该存储器单元一次；以及

9*B*）该组合电路还包括：

9*B*1）能够与存储器单元及优化增益确定器耦合操作的一个滤波器加权确定器，用于确定滤波器的多个系数；

9*B*2）能够与存储器单元及滤波器加权确定器耦合操作的多个多路复用器，用于产生多个加权的解压缩的象素强度；以及

9B3）能够与该多个多路复用器耦合操作的一个加法器，用于产生后处理象素强度。

10. 抑制视频压缩系统，其中包括一个用于抑制蚊形及分块人工形成物的后处理滤波器，该系统包括：

A）能够与接收视频序列耦合操作的一个视频编码单元，用于去除包含在视频序列中的冗余信息以产生压缩的视频流；

B）能够与视频编码单元耦合操作的一个信道编码单元，用于对压缩的视频流进行编码以产生设置在一个信道上的编码的压缩的视频流；

C）能够通过该信道到信道编码单元耦合操作的一个信道解码单元，用于对编码的压缩的视频流进行解码以产生压缩的视频流；

D）能够与信道解码单元耦合操作的一个视频解码单元，用于产生一个解压缩的视频序列；以及

E）能够与该视频解码单元耦合操作的一个后处理单元，用于使用后处理滤波器，其中该后处理滤波器从解压缩的视频序列去除蚊形人工形成物和分块人工形成物以提供一个后处理的视频序列。

说　明　书

用于抑制人工形成物的方法，后处理滤波器与视频压缩系统

本发明一般涉及视频压缩系统，并特别涉及抑制视频压缩系统中蚊形和分块人工形成物。

在视频压缩系统中，一个视频序列中表示一个信号帧所需要的位数可利用信号帧所具有的与前面的帧的相似性而被降低。已经使用了诸如混合块运动补偿离散余弦变换（DCT）编码方案等方法而达到了很高的压缩比，但是这些方法也引入了严重降低解码序列视频质量的人工形成物。

所引起的两种人工形成物是蚊形人工形成物与分块人工形成物。蚊形人工形成物定义为围绕在解压缩视频内运动的对象出现的瞬时不稳定脉冲。这些人工形成物是由于预测误差信号的不精确量化的结果。预测误差信号中所包含的大部分能量是由于运动估计者不能区分不同地运动着的对象的结果。例如，在电视会议应用中，主题一般是与固定的背景相对的。由于运动估计者试图在暂时邻接的帧之间匹配象素块，落在这些块内的运动着的对象与固定的背景之间的边界就不能被检测到。这导致或者部分的边界被认为是在运动着、或者运动着的对象的部分被认为是固定的这样的情形。这些预测误差不精确量化的结果造成随时间变化并象要群集在运动对象周围类似于蚊子样的脉冲型人工形成物。

分块人工形成物定义为向解码的视频序列引入人工块边界。这些人工形成物是由于划分预测误差信号为块，以及量化组合的结果。即，由于在空间扩展和频率扩展之间存在一种相反的关系，这类似于傅立叶分析中时间和频率扩展之间存在的相反的关系，故出现在 DCT 域中的量化误差分布到相应的空间块。而且，由于每一块是被分开量化的，故误差大部分在块边界看到。

由于这些人工形成物是在特定的位置出现的，并且不均匀地出现在视频序列各处，设计用来减少这两种人工形成物的后处理滤波器必须在空间上适配。空间适配的滤波器已经用于抑制这些人工形成物，但是它们依赖于局部信号及噪声功率估计而改变其响应。这种设计存在若干问题。例如，基于量化步长大小的噪声功率估计对于在解压缩视频内的蚊形和分块人工形成物就不是一个可靠的指标。出现解压缩视频的过光滑或者模糊是由于对解压缩视频信号对噪声的比率（SNR）不精确的估计。分块人工形成物会因在解压缩视频中不正确地将块边界划分为对象边界而加重。解压缩视频上可觉察的视觉质量的这些问题中的任何一个，其效果都是灾难性的。

进一步的考虑是压缩视频用户要求解码器要能够重新设定视频显示窗口为任意指定大小。为了向用户提供这种能力，要求解码器进行关于去除编码的人工形成物、涉及数据内插或者细分的额外的后处理操作。一个重要的要求是，这种重新设定尺寸的操作是计算上有效的并且提供视觉上有感染力的尺寸再设定序列。这就是说，不应当由于尺寸再设定操作的结果而向视频序列引入人工形成物。此外，编码的人工形成物不应当由于尺寸再设定操作而在解码序列中变得更明显。

于是，需要一种与局部信号对噪声功率比率估计无关的能够抑制蚊形和分块人工形成物的方法，后处理滤波器，以及视频压缩系统，同时还提供再设定解压缩视频尺寸的能力。

图 1 是由根据本发明的抑制蚊形和分块人工形成物的后处理滤波器所执行的步骤的流程图。

图 2 是使用根据本发明的抑制蚊形和分块人工形成物的后处理滤波器的视频压缩系统的框图。

图 3 是根据本发明的抑制蚊形和分块人工形成物的后处理滤波器的框图。

图 4 是根据本发明使用预测器和更新器的抑制蚊形和分块人工形成物的后处理滤波器的框图。

图 5 是根据本发明抑制蚊形和分块人工形成物并进行尺寸再设定的后处理滤波器的框图。

本发明使得视频压缩系统中的后处理滤波器能够抑制蚊形和分块人工形成物并且是与局部信号对噪声功率比率估计无关的。基于象素的局部邻域对信号的方差进行估计。象素的局部邻域包括多个前面后处理过的的象素强度及多个原有象素强度。基于信号方差，预定的噪声方差，以及预定的自相关系数计算优化的滤波器增益。应用预定的自相关系数，象素的局部邻域，以及优化的滤波器增益确定后处理的象素强度。而且，如果需要，可应用预定的自相关系数，象素的局部邻域，优化的滤波器增益以及多个内插系数确定后处理的尺寸再设定的象素强度。这消除了对于局部信号对噪声功率比率估计的依赖性，同时向用户提供了对于解压缩视频尺寸再设定和后处理的能力。

参照图1—5对本发明进行更为充分的描述。标号100的图1是由根据本发明的抑制蚊形和分块人工形成物的后处理滤波器所执行的步骤的流程图。首先，基于解压缩视频序列在存储器单元中生成象素的局部邻域(101)。然后基于象素的局部邻域估计一个信号方差(102)。然后，应用该信号方差，预定的噪声方差，以及预定的自相关系数计算优化的滤波器增益(104)。最后，基于预定的自相关系数，象素的局部邻域，以及优化的滤波器增益确定后处理的象素强度(106)。

当需要再设定解压缩的视频窗口尺寸时，后处理的象素强度是基于多个预定的内插系数的再设定尺寸的后处理的象素强度。这是通过结合尺寸再设定操作与上述的后处理滤波器而实现的。后处理的象素从局部邻域除去而以解压缩象素替代。以上所述的参数，基于象素局部邻域的信号方差和基于信号方差的优化滤波器增益，预定的噪声方差，以及预定的自相关系数与多个内插系数结合而形成这一结合的尺寸再设定后处理滤波器的系数。最后，尺寸再设定后处理象素强度基于预定的自相关系数，象素的局部邻域，优化的滤波器增益，以及多个内插系数而确定。

该信号方差可使用形式为

$$\hat{\sigma}_s^2(n)=\frac{1}{2M+1}[\sum_{i=-M}^{0}(x(n-i))^2+\sum_{i=1}^{M}(\hat{y}(n-i))^2]-\hat{m}_s^2(n),$$

的方程式估计，其中

$$\hat{m}_s(n)=\frac{1}{2M+1}\left[\sum_{i=-M}^{0}x(n-i)+\sum_{i=1}^{M}\hat{y}(n-i)\right]$$

是象素的$(2M+1)$局部邻域中的平均强度值，$\hat{y}(n)$是象素的局部邻域中先前后处理的的象素强度，而$x(n)$是象素的局部邻域中解压缩象素强度。优化的滤波器增益可应用形式为：

$$G(n)=\frac{\alpha \bullet \sigma_s^2(n)}{\sigma_n^2+\alpha \bullet \sigma_s^2(n)}$$

的方程式计算，其中

$$\alpha=1-\frac{2 \bullet \zeta^2}{1+\zeta^2}$$

以及σ_n^2是噪声方差，$\sigma_n^2(n)$是在象素位置n的信号方差，而ζ是预定的自相关系数。噪声方差的典型区域是$0<\sigma_n^2<10$。自相关系数的典型区域是$0.9<\zeta<1.0$。后处理象素强度可基于象素强度的估计。象素强度的估计是预定的自相关系数、相对于预定的扫描方向下一个象素的解压缩象素的强度、以及紧靠前面的一个后处理象素强度的函数。象素强度的估计可应用形式为：

$$\hat{y}_a(n)=\frac{\zeta}{1+\zeta^2} \bullet \left[\hat{y}_b(n-1)+x(n+1)\right]$$

的方程式确定，其中$\hat{y}_a(n)$为一个象素强度的估计，ζ是预定的自相关系数，$x(n+1)$是相对于预定的扫描方向下一个象素的解压缩象素的强度，$\hat{y}_b(n-1)$是紧靠前面的一个后处理象素强度。后处理象素

强度可应用形式为：

$$\hat{y}_b(n)=\hat{y}_a(n)+G(n)\bullet\left[x(n)-\hat{y}_a(n)\right]$$

的方程式确定，其中 $G(n)$ 为优化滤波器增益，$\hat{y}_a(n)$ 是象素局部邻域中象素强度的估计，$x(n)$ 是该象素局部邻域中解压缩的象素强度。

对于形式为：

$$y(m)=C_0\cdot x(n)+C_1\cdot x(n+1)+C_{-1}\cdot x(n-1)+C_2\cdot x(n+2)+C_{-2}\cdot x(n-2)+\ldots\ldots,$$

所产生的尺寸再设定滤波器，其中 $y(m)$ 是位于 $x(n)$ 与 $x(n+1)$ 之间尺寸再设定解压缩象素，而 C_i 是预定的内插系数，该系数取决于内插或者细分方法，以下方程式用于确定尺寸再设定后处理象素：

$$\begin{aligned}\hat{y}_b(m)=&[C_{-1}\cdot g(n)+C_0\cdot(1-g(n))\cdot \mathrm{B}]x(n-1)+[C_0\cdot g(n)+(1-g(n))\cdot(C_{-1}\cdot\zeta+C_1\cdot \mathrm{B}]x(n)\\&+[C_1\cdot g(n)+(1-g(n))\cdot(C_2\cdot\zeta+C_0\cdot \mathrm{B}]x(n+1)+[C_2\cdot g(n)+C_1\cdot(1-g(n))\cdot \mathrm{B}]x(n+2)\\&+C_{-2}\cdot x(n-2)+C_3\cdot x(n+3)+C_{-3}\cdot x(n-3)+\ldots\end{aligned}$$

其中

$$\mathrm{B}=\frac{\zeta}{1+\zeta^2},$$

而 $\hat{y}_b(m)$ 是位于 $x(n)$ 与 $x(n+1)$ 之间的尺寸再设定后处理象素。

标号为200的图2是使用根据本发明的抑制蚊形和分块人工形成物的后处理滤波器的视频压缩系统的框图。该视频压缩系统(200)包括一个视频编码单元(204)，一个信道编码单元(208)，一个信道(210)，一个信道解码单元(212)，一个视频解码单元(214)，和一个包含后处理滤波器(222)的后处理单元(218)。

包含在输入视频序列(202)中的冗余信息由产生压缩视频流(206)的视频编码单元(204)消除。然后被压缩的视频流(206)被信道编码单元(208)编码用于通过信道(210)有效传输到对应的信道解码单元(212)。信道解码单元(212)对所接收的信息进行解码并提供向视频解码单元(214)提供压缩视频流(213)。基于压缩视频流(213)，视频解码单元(214)产生解压缩视频序列(216)。这一解压缩视频序列(216)是向后处理单元(218)的输入，其中使用后处理滤波器(222)从产生后处理视频序列(220)的解压缩视频中消除蚊形和分块人工形成物。

标号为300的图3是根据本发明的抑制蚊形和分块人工形成物的后处理滤波器的框图。后处理滤波器(222)包括一个存储器单元(302)，一个信号方差确定器(310)，一个优化的滤波器增益确定器(314)，以及组合电路(318)。存储器单元(302)接收和存储一个预定的噪声方差(304)，一个预定的自相关系数(306)，和一个解压缩视频序列(216)。该解压缩视频序列(216)被存储为象素的一个局部邻域(308)。

存储在存储器单元(302)中的预定的噪声方差一般是1到10之间的一个数值。信号方差确定器(310)基于象素的局部邻域(308)估算信号方差(312)。优化滤波器增益确定器(314)基于信号方差

(312)，预定的噪声方差(304)，和预定的自相关系数(306)确定优化滤波器增益(316)。组合电路(318)基于预定的自相关系数(306)，象素的局部邻域(308)，以及优化滤波器增益(316)确定后处理的象素强度(320)。组合电路(318)可作为一个无限脉冲响应(*IIR*)滤波器或者一个有限脉冲响应(*FIR*)滤波器而实现。

标号400的图4是根据本发明使用预测器(402)和更新器(404)的抑制蚊形和分块人工形成物的后处理滤波器的框图。该后处理滤波器包括一个存储器单元(302)，一个信号方差确定器(310)，一个优化滤波器增益确定器(314)，和一个组合电路(318)。存储器单元(302)接收和存储预定的噪声方差(304)，预定的自相关系数(306)，及解压缩的视频序列(216)。解压缩的视频序列(216)是作为象素的一个局部邻域(308)而存储的。应当注意，图4反映了当$M=2$时对应于由*X´s*与*O´s*所指示的五个象素元素所组成的象素局部邻域的情形。*M*的选择在执行时进行并且取决于所需要的光滑程度：*M*越大施加到解压缩视频的滤除越多。假定左到右通过图象帧的每一行，*X´s*和*O´s*指示那个象素是先前后处理的强度的象素，*X´s*，以及那个是解压缩强度的象素，*O´s*。其它方法也可用于通过存储在存储器中的解压缩图象帧的运动，诸如从右到左通过每一行，从上到下或者从下到上通过每一列。

图4中所示的框图的一般操作与图3中所示的框图的操作相同，并对于组合电路(318)的实现更为详细示出。组合电路(318)包括一个预测器(402)和一个更新器(404)。预测器(402)预测当前象素的强度。更新器(404)确定后处理象素强度。

预测器(402)使用象素的局部邻域(308)的一维非表因(*non*

—*causal*）自回归模型。一个加法器(408)使得第一个先前的后处理象素($\hat{y}_b(n-1)$)与第一个解压缩象素($x(n+1)$)越过在考虑之下的象素($x(n)$)相加。这一运算所得的结果是在这两个象素位置的当前的强度水平之和(409)。然后由一阶预测系数确定器(406)所提供的系数(407)用于乘上(410)强度的和(409)。结果这得到在象素位置 n 处的原始强度的估计值(412)。然后估计值(412)与解压缩象素强度($x(n)$)比较(414)。然后这一比较的输出(415)乘以(416)优化滤波器增益确定器(314)的输出(316)。这一数值的输出(417)使用加法器(418)与强度估计值(412)求和，其结果为后处理的象素强度(320)。

标号为 500 的图 5 是根据本发明抑制蚊形和分块人工形成物并进行尺寸再设定的后处理滤波器的框图。该后处理滤波器包括一个存储器单元(302)，一个信号方差确定器(310)，一个优化滤波器增益确定器(314)，以及一个组合电路(318)。存储器单元(302)接收和存储预定的噪声方差(304)，预定的自相关系数(306)，和解压缩视频序列(216)。解压缩的视频序列(216)是作为象素的局部邻域(308)存储的。如图 4 中那样，图 5 反映了当 $M=2$ 时对应于由 O′s 所表示的五个象素成分所组成的象素局部邻域的情形。M 的选择是在执行时作出的，并取决于所希望的光滑程度和所选的尺寸再设定滤波器的类型：M 越大施加到解压缩视频的滤除越多。O′s 指示哪些解压缩强度象素要用于局部邻域中。

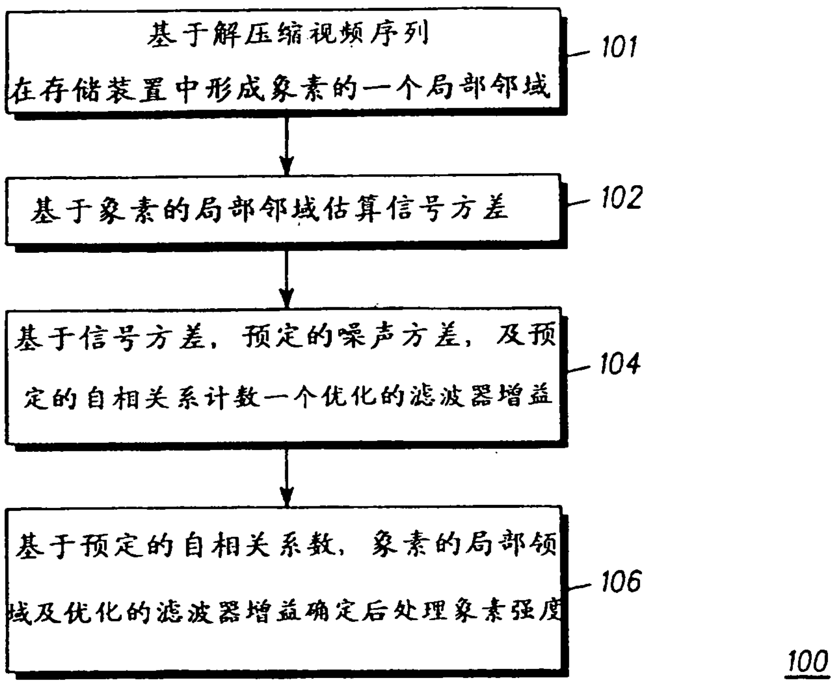
存储器单元(302)还存储预定的多个内插系数(501)。组合电路(318)一般包括一个尺寸再设定滤波器。该组合电路存取存储器(302)一次。该组合电路(318)包括一个滤波器加权确定器(502)，多

个多路复用器(504),和一个加法器(506)。该滤波器加权确定器(502)确定多个滤波器系数(508)。多个多路复用器(504)产生多个加权的解压缩象素强度(510)。加法器提供了后处理象素强度(320)。

该后处理滤波器(400到500)可以使用诸如 *MOTOROLA DSP*56002 的一个数字信号处理器(*DSP*)或者一个专用集成电路(*ASIC*)和一个随机存取存储器(*RAM*)单元实现。存储器单元(302),可以是 *RAM* 单元。信号方差确定器(310),优化滤波器增益确定器(314),以及预测器单元(402)可使用该 *ASIC* 实现。更新器可使用该 *ASIC* 之中的两个加法器(414及418)和一个多路复用器(416)实现。

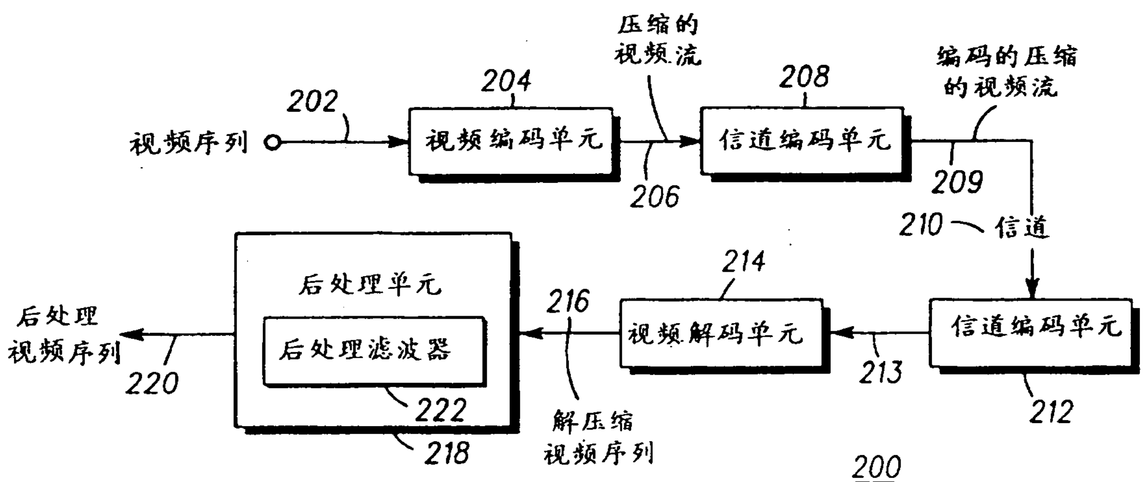
本发明提供了用于从解压缩的视频序列去除蚊形及分块人工形成物的方法和装置。利用这种方法和装置,可以对解压缩的视频进行尺寸再设定,同时去除这些人工形成物。这一方法和装置具有胜过当前的后处理滤波器的以下优点:它与信号对噪声的计算结果这一人工形成物位置的不良的指标无关;它能够通过信号对存储器的访问进行尺寸再设定及后处理;并获得较高的可觉察的收视质量。

虽然以上对示例性的实施方式进行了说明,显然对于业内人士来说在不背离本发明的情形下可作出各种变化和改形。于是,意思是说所有那些变化和改形均包括在所附权利要求中定义的本发明的精神和范围之内。



基于解压缩视频序列在存储装置中形成象素的一个局部邻域
101
基于象素的局部邻域估算信号方差
102
基于信号方差，预定的噪声方差，及预定的自相关系计数一个优化的滤波器增益
104
基于预定的自相关系数，象素的局部领域及优化的滤波器增益确定后处理象素强度
106
100

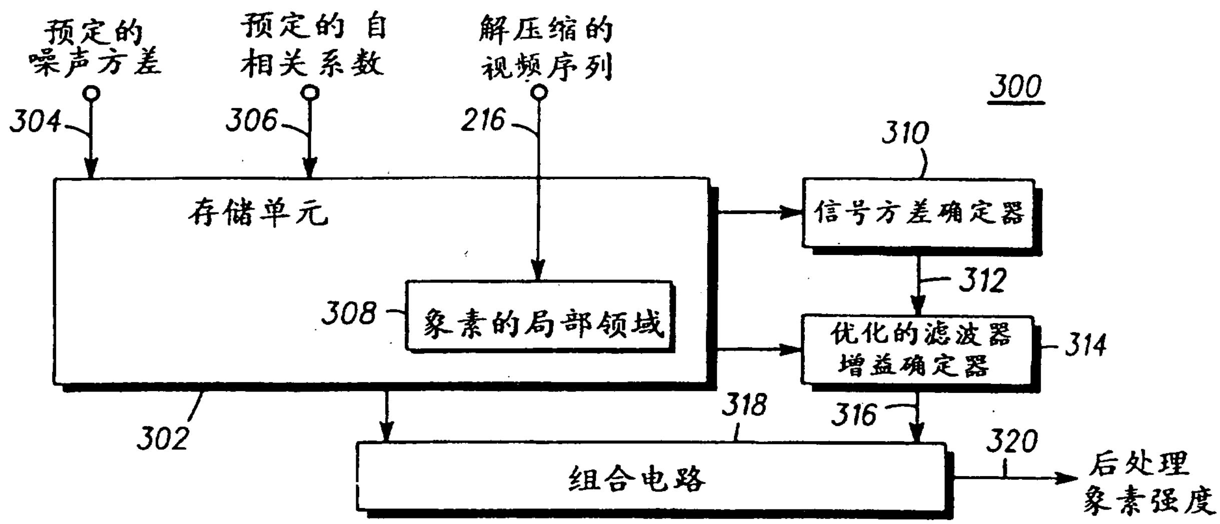
图 1



视频序列
202
视频编码单元
204
压缩的视频流
206
信道编码单元
208
编码的压缩的视频流
209
210 信道
信道编码单元
212
213
视频解码单元
214
216
解压缩视频序列
后处理单元
后处理滤波器
222
218
220
后处理视频序列
200

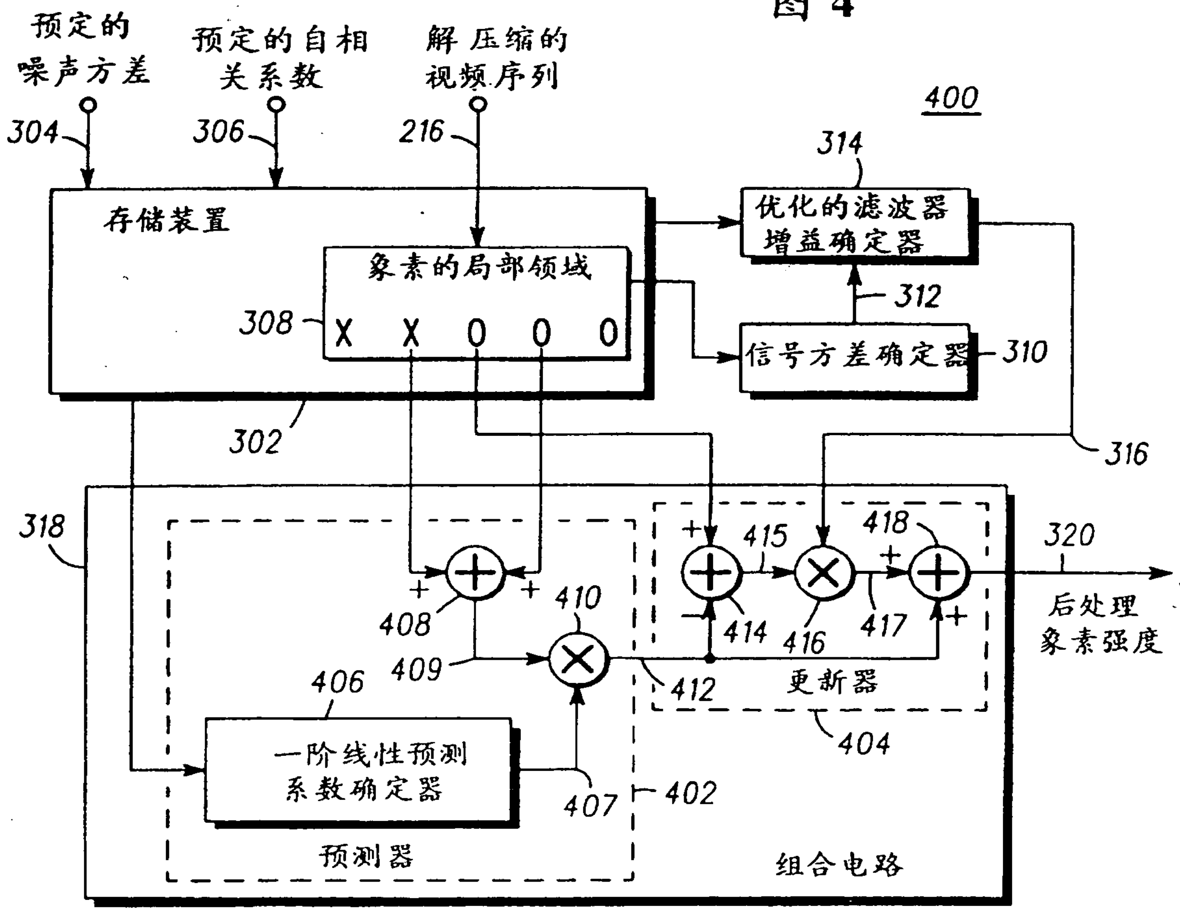
图 2

图 3



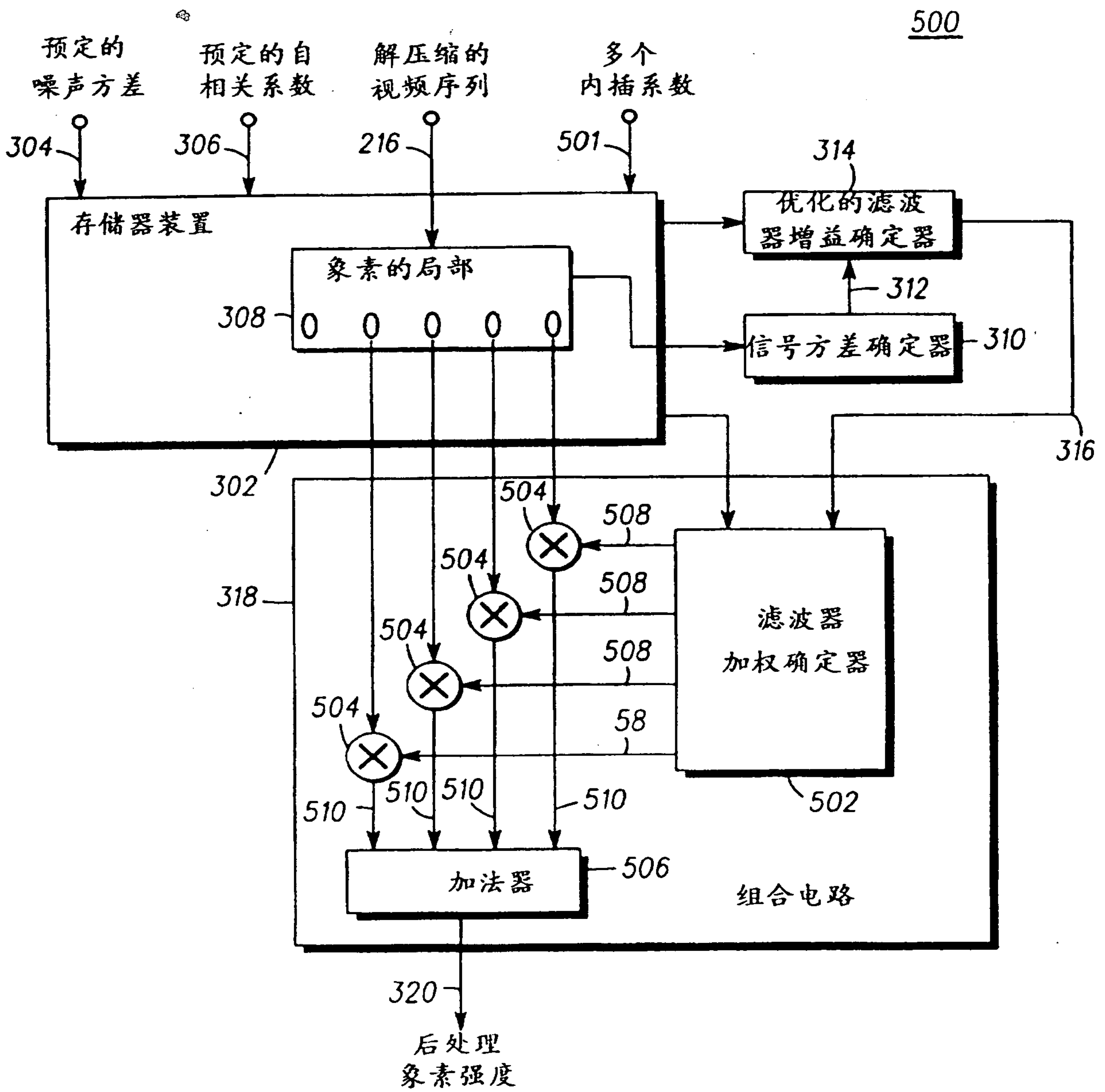
预定的
噪声方差
预定的自
相关系数
解压缩的
视频序列
300
304
306
216
310
存储单元
信号方差确定器
312
308
象素的局部领域
优化的滤波器
增益确定器
314
318
316
302
320
组合电路
后处理
象素强度

图 4



预定的
噪声方差
预定的自相
关系数
解压缩的
视频序列
400
304
306
216
314
存储装置
优化的滤波器
增益确定器
象素的局部领域
312
308
X X 0 0 0
信号方差确定器
310
302
316
318
415
418
320
408
410
414
416
417
后处理
象素强度
409
406
412
更新器
404
一阶线性预测
系数确定器
407
402
预测器
组合电路

图 5

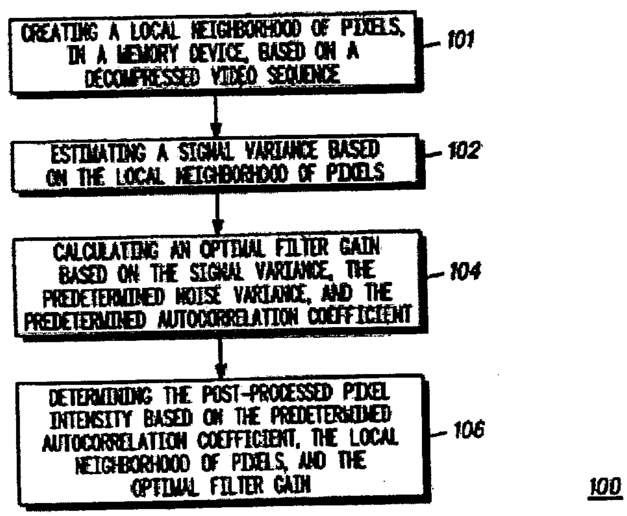

500
预定的
噪声方差
预定的自
相关系数
解压缩的
视频序列
多个
内插系数
304
306
216
501
314
存储器装置
优化的滤波
器增益确定器
象素的局部
312
308
信号方差确定器
310
302
316
504
508
318
滤波器
加权确定器
58
510
502
加法器
506
组合电路
320
后处理
象素强度

Method, post-processing filter, and video compression system for suppressing mosquito and blocking artifacts

Patent number: CN1138401
Publication date: 1996-12-18
Inventor: BRAILEAN JAMES CHARLES (US)
Applicant: MOTOROLA INC (US)
Classification:
- international: *G06T5/20; H04N7/26; H04N7/30; H04N5/21;* **G06T5/20; H04N7/26; H04N7/30**; H04N5/21; (IPC1-7): H04N1/415; G06T5/00
- european: G06T5/20; H04N7/26P4; H04N7/30P
Application number: CN19950191135 19950831
Priority number(s): US19940334718 19941104

Abstract not available for CN1138401
Abstract of corresponding document: **US5802218**

The visibility of mosquito and blocking artifacts in a video compression system (200) are effectively suppressed by calculating a post-processed pixel intensity. A post-processed pixel intensity is determined (100) using a predetermined noise variance, a predetermined autocorrelation coefficient, and a local neighborhood of pixels that includes a plurality of previously post-processed pixel intensities and a plurality of original pixel intensities. This eliminates the dependency on local signal and noise power estimates.



Method, post-processing filter, and video compression system for suppressing mosquito and blocking artifacts

Claims of corresponding document: **US5802218**

I claim:

1. A method for obtaining a post-processed pixel intensity, for a frame in a decompressed video sequence residing in memory, using a filter which suppresses a visibility of a mosquito artifact and a blocking artifact based on a predetermined noise variance and a predetermined autocorrelation coefficient, comprising the steps of:
A) creating a local neighborhood of pixels, in a memory unit, based on the decompressed video sequence, wherein the local neighborhood of pixels includes a plurality of previously post-processed pixel intensities and a plurality of decompressed pixel intensities;
B) estimating a signal variance based on the local neighborhood of pixels;
C) calculating an optimal filter gain based on the signal variance, the predetermined noise variance, and the predetermined autocorrelation coefficient; and
D) determining, using a summer, the post-processed pixel intensity based on the predetermined autocorrelation coefficient, the local neighborhood of pixels, and the optimal filter gain,
wherein the post-processed pixel intensity is further based on an estimate of an intensity of a selected pixel,
wherein the estimate of the intensity of the selected pixel is a function of the predetermined autocorrelation coefficient, a decompressed pixel intensity of a next pixel relative to a predetermined scanning, and direction, and an immediately previously post-processed pixel intensity,
wherein the estimate of the intensity of the selected pixel is determined by using an equation of the form: ##EQU5## where ya (n) is the estimate of the intensity of the selected pixel, .zeta. is the predetermined autocorrelation coefficient, x(n+1) is a decompressed pixel intensity of a next pixel relative to a predetermined scanning direction, and yb (n-1) is an immediately previously post-processed pixel intensity.

2. The method of claim 1, wherein the signal variance is estimated by using an equation of the form: ##EQU6## is an average intensity value in the local neighborhood of pixels, (2M+1) a neighborhood size, and is y(n) is a previously post-processed pixel intensity in the local neighborhood of pixels.

3. The method of claim 1, wherein the optimal filter gain is calculated by using an equation of the form: ##EQU7## and where .sigma.@2n is the predetermined noise variance, .sigma.@2s is the signal variance, and .zeta. is the predetermined autocorrelation coefficient.

4. The method of claim 1, wherein the post-processed pixel intensity is a resized post-processed pixel intensity based on a plurality of predetermined interpolation coefficients.

5. The method of claim 4, wherein memory is accessed one time for determining the resized post-processed pixel intensity.

6. A method for obtaining a post-processed pixel intensity, for a frame in a decompressed video sequence residing in memory, using a filter which suppresses a visibility of a mosquito artifact and a blocking artifact based on a predetermined noise variance and a predetermined autocorrelation coefficient, comprising the steps of:
A) creating a local neighborhood of pixels, in a memory unit, based on the decompressed video sequence, wherein the local neighborhood of pixels includes a plurality of previously post-processed pixel intensities and a plurality of decompressed pixel intensities;
B) estimating a signal variance based on the local neighborhood of pixels;
C) calculating an optimal filter gain based on the signal variance, the predetermined noise variance, and the predetermined autocorrelation coefficient; and
D) determining, using a summer, the post-processed pixel intensity based on the predetermined autocorrelation coefficient, the local neighborhood of pixels, and the optimal filter gain, and
wherein the post-processed pixel intensity is further based on an estimate of an intensity of a selected pixel,
wherein the post-processed pixel intensity is determined by using an equation of the form:

yb (n)=ya (n)+G(n).cndot.[x(n)-ya (n)]

where G(n) is the optimal filter gain, ya (n) is the estimate of the intensity of a pixel, and x(n) is a decompressed pixel intensity.

7. A video compression system which includes a post-processing filter for suppressing a visibility of a mosquito artifact and a blocking artifact, the system comprising:
A) a video encoding unit, operably coupled to receive a video sequence, for removing redundant information contained within the video sequence to generate a compressed video stream;
B) a channel encoding unit, operably coupled to the video encoding unit, for encoding the compressed video stream to generate an encoded compressed video stream which is set over a channel;
C) a channel decoding unit, operably coupled through the channel to the channel encoding unit, for decoding the encoded compressed video stream to generate the compressed video stream;
D) a video decoding unit, operably coupled to the channel decoding unit, for generating a decompressed video sequence; and
E) a post-processing unit, operably coupled to the video decoding unit, for utilizing a post-processing filter, wherein the post-processing filter removes the mosquito artifact and blocking artifact from the decompressed video sequence based on a predetermined noise variance and a predetermined autocorrelation coefficient to provide a post-processed video sequence, wherein the predetermined noise variance is a value between 1 and 10, by creating a local neighborhood of pixels in a memory unit, based on the decompressed video sequence, wherein the local neighborhood of pixels includes a plurality of previously post-processed pixel intensities and a plurality of decompressed pixel intensities, estimating a signal variance based on the local neighborhood of pixels; calculating an optimal filter gain based on the signal variance, the predetermined noise variance, and the predetermined autocorrelation coefficient; and determining, using a summer, the

post-processed pixel intensity based on the predetermined autocorrelation coefficient, the local neighborhood of pixels, and the optimal filter gain, wherein the post-processed pixel intensity is further based on an estimate of an intensity of a selected pixel, wherein the estimate of the intensity of the selected pixel is a function of the predetermined autocorrelation coefficient, a decompressed pixel intensity of a next pixel relative to a predetermined scanning, and direction, and an immediately previously post-processed pixel intensity, and wherein the estimate of the intensity of the selected pixel is determined by using an equation of the form: ##EQU8## where ya (n) is the estimate of the intensity of the selected pixel, .zeta. is the predetermined autocorrelation coefficient, x(n+1) is a decompressed pixel intensity of a next pixel relative to a predetermined scanning direction, and yb (n-1) is an immediately previously post-processed pixel intensity.

Method, post-processing filter, and video compression system for suppressing mosquito and blocking artifacts

Description of corresponding document: US5802218

FIELD OF THE INVENTION

The present invention relates generally to video compression systems, and more particularly to suppression of mosquito and blocking artifacts in a video compression system.

BACKGROUND

In video compression systems, the number of bits required to represent a single frame in a video sequence can be reduced by taking advantage of the similarities that the single frame has with previous frames. Methods such as a hybrid block motion compensated discrete cosine transform, DCT, encoding scheme have been used to achieve very high compression ratios, but these methods also introduce artifacts which severely degrade the visual quality of the decoded sequence.

The two types of artifacts introduced are mosquito artifacts and blocking artifacts. Mosquito artifacts are defined as temporally nonstationary impulses that appear around objects which are moving within the decompressed video. These artifacts result from the coarse quantization of the prediction error signal. The majority of the energy contained in a prediction error signal is the result of the motion estimator's inability to distinguish between differently moving objects. For instance, in video conferencing applications the subject is generally against a stationary background. Since the motion estimator tries to match blocks of pixels between temporally adjacent frames, the boundaries between moving objects and stationary background that fall within these blocks cannot be detected. This leads to the situation where either part of the background is assumed to be moving, or part of the moving object is assumed to be stationary. Coarsely quantizing these prediction errors results in impulsive artifacts that change over time and tend to swarm around the moving object, similar in manner to a mosquito.

Blocking artifacts are defined as the introduction of artificial block boundaries into the decoded video sequence. These artifacts are due to the combination of dividing the prediction error signal into blocks, as well as quantization. That is, since there exists an inverse relationship between spatial extent and frequency extent analogous to the inverse relationship that exists between time and frequency extent in Fourier analysis, the quantization errors that occur in the DCT domain are smeared across the corresponding spatial block. Furthermore, since each block is quantized separately, the errors are most visible at the block boundaries.

Since both of these artifacts occur at specific locations and not uniformly throughout the video sequence, post-processing filters, designed to reduce both artifacts, must be spatially adaptive. Spatially adaptive filters have been used to suppress these artifacts, but they rely on local signal and noise power estimates to alter their responses. Several problems exist with this design. For instance, estimation of the noise power based on the quantization step size is not a reliable

indicator as to the spatial location of mosquito and blocking artifacts within the decompressed video. Over-s oothing or blurring of the decomp ssed video occurs due to inaccurate estimates of the decompressed video's signal-to-noise ratio, SNR. While enhancement of the blocking artifacts can result from the incorrect classification of block boundaries as object edges within the decompressed video. The effect of any one of these problems on the perceived visual quality of the decompressed video is disastrous.

A further consideration is that users of compressed video demand that the decoder be able to resize the video display window to any specified size. To provide a user with this capability, requires the decoder to perform an additional, with regards to removing coding artifacts, post-processing operation involving data interpolation or decimation. An important requirement is that this resize operation must be computationally efficient and provide a resize sequence that is visually appealing. That is, artifacts should not be introduced into the video sequence as a result of the resizing operation. Furthermore, coding artifacts should not become more visible in the decoded sequence due to the resizing operation.

Thus, there is a need for a method, post-processing filter, and video compression system which suppress mosquito and blocking artifacts independent of local signal to noise power ratio estimates, while also providing the ability to resize the decompressed video.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a flow chart of steps implemented by a post-processing filter which suppresses mosquito and blocking artifacts in accordance with the present invention.

FIG. 2 is a block diagram of a video compression system that uses a post-processing filter which suppresses mosquito and blocking artifacts in accordance with the present invention.

FIG. 3 is a block diagram of a post-processing filter which suppresses mosquito and blocking artifacts in accordance with the present invention.

FIG. 4 is a block diagram of a post-processing filter which suppresses mosquito and blocking artifacts, using a predictor and an updater, in accordance with the present invention.

FIG. 5 is a diagram of a post-processing filter which suppresses mosquito and blocking artifacts and performs resizing in accordance with the present invention.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

The present invention allows a post-processing filter in a video compression system to suppress mosquito and blocking artifacts and to be independent of local signal and noise power estimates. A signal variance is estimated based on a local neighborhood of pixels. The local neighborhood of pixels includes a plurality of previously post-processed pixel intensities and a plurality of original pixel intensities. An optimal filter gain is calculated based on the signal variance, a predetermined noise variance, and a predetermined autocorrelation coefficient. A

post-processed pixel intensity is determined using the predetermined autocorrelation coefficient, the local neighborhood of pixels, and the optimal filter gain. Also, if desired, a post-processed resized pixel intensity is determined using the predetermined autocorrelation coefficient, the local neighborhood of pixels, the optimal filter gain, and plurality of interpolation coefficients. This eliminates the dependency on local signal to noise power ratio estimates, while providing the user with the ability to resize and post-process the decompressed video.

The present invention is more fully described with reference to FIGS. 1-5. FIG. 1, numeral 100, is a flow chart of steps implemented by a post-processing filter which suppresses mosquito and blocking artifacts in accordance with the present invention. First, a local neighborhood of pixels is created, in a memory unit, based on a decompressed video sequence (101). Then, a signal variance is estimated based on a local neighborhood of pixels (102). An optimal filter gain is then calculated using the signal variance, a predetermined noise variance, and a predetermined autocorrelation coefficient (104). Finally, a post-processed pixel intensity is determined based on the predetermined autocorrelation coefficient, the local neighborhood of pixels, and the optimal filter gain (106).

When resizing of the decompressed video window is required, the post-processed pixel intensity is a resized post-processed pixel intensity based on a plurality of predetermined interpolation coefficients. This is accomplished by combining the resizing operation with the post-processing filter described above. The post-processed pixels are removed from the local neighborhood and replaced with decompressed pixels. The parameters described above, the signal variance based on a local neighborhood of pixels and the optimal filter gain based on the signal variance, a predetermined noise variance, and a predetermined autocorrelation coefficient, are combined with a plurality of interpolation coefficients to form the coefficients of this combined resize post-processed filter. Finally, a resized post-processed pixel intensity is determined based on the predetermined autocorrelation coefficient, the local neighborhood of pixels, the optimal filter gain, and plurality of interpolation coefficients.

The signal variance may be estimated by using an equation of the form: ##EQU1## is the average intensity value in the (2M+1) local neighborhood of pixels, y(n) is a previously post-processed pixel intensity in the local neighborhood of pixels, and x (n) is a decompressed pixel intensity in the local neighborhood of pixels. The optimal filter gain may be calculated by using an equation of the form: ##EQU2## and where .sigma.@2n is the noise variance, .sigma.@2s (n) is the signal variance at pixel location n, and .zeta. is the predetermined autocorrelation coefficient. A typical range for the noise variance is 0<.sigma.@2n <10. A typical range for the autocorrelation coefficient is 0.9<.zeta.<1.0. The post-processed pixel intensity may be based on an estimate of the intensity of a pixel. The estimate of the intensity of a pixel is a function of the predetermined autocorrelation coefficient, a decompressed pixel intensity of a next pixel relative to a predetermined scanning direction, and an immediately previously post-processed pixel intensity. The estimate of the intensity of a pixel may be determined by using an equation of the form: ##EQU3## where ya (n) is the estimate of the intensity of a pixel, .zeta. is the predetermined autocorrelation coefficient, x(n+1) is the decompressed pixel intensity of a next pixel relative to a predetermined scanning direction, and yb (n-1) is an immediately previously post-processed pixel intensity. The post-processed pixel intensity may determined by using an equation of the form:

yb (n)=ya (n)+G(n).cndot.[x( )-ya (n)]

where G(n) is the optimal filter gain, ya (n) is the estimate of the intensity of a pixel in the local neighborhood of pixels, and x(n) is the decompressed pixel intensity in the local neighborhood of pixels.

For the generalized resizing filter of the form:

y(m)=C0 .multidot.x(n)+C1 .multidot.x(n+1)+C-1 .multidot.x(n-1)+C2 .multidot.x (n+2)+C-2 .multidot.x(n-2)+,

where y(m) is the resized decompressed pixel located between x(n) and x(n+1) and Ci is a predetermined interpolation coefficient which is dependent on the interpolation or decimation method, the following equation is used to determine the resized post-processed pixel: ##EQU4## where and yb (m) is the resized post-processed pixel located between x(n) and x(n+1).

FIG. 2, numeral 200, is a block diagram of a video compression system that uses a post-processing filter which suppresses mosquito and blocking artifacts in accordance with the present invention. The video compression system (200) includes a video encoding unit (204), a channel encoding unit (208), a channel (210), a channel decoding unit (212), a video decoding unit (214), and a post-processing unit (218) which includes a post-processing filter (222).

Redundant information contained within the input video sequence (202) is removed by the video encoding unit (204) generating the compressed video stream (206). The compressed video stream (206) is then encoded by the channel encoding unit (208) for efficient transmission over the channel (210) to the corresponding channel decoding unit (212). The channel decoding unit (212) decodes the received information and provides the video decoding unit (214) with the compressed video stream (213). Based on the compressed video stream (213), the video decoding unit (214) generates the decompressed video sequence (216). This decompressed video sequence (216) is the input to the post-processing unit (218), where the post-processing filter (222) is employed to remove the mosquito and blocking artifacts from the decompressed video resulting in the post-processed video sequence (220).

FIG. 3, numeral 300, is a block diagram of a post-processing filter which suppresses mosquito and blocking artifacts in accordance with the present invention. The post-processing filter (222) comprises a memory unit (302), a signal variance determiner (310), an optimal filter gain determiner (314), and a combining circuit (318). The memory unit (302) receives and stores a predetermined noise variance (304), a predetermined autocorrelation coefficient (306), and a decompressed video sequence (216). The decompressed video sequence (216) is stored as a local neighborhood of pixels (308).

The predetermined noise variance stored in the memory unit (302) is typically a value between 1 and 10. The signal variance determiner (310) estimates the signal variance (312) based on the local neighborhood of pixels (308). The optimal filter gain determiner (314) determines an optimal filter gain (316) based on the signal variance (312), the predetermined noise variance (304), and the predetermined autocorrelation coefficient (306). The combining circuit (318) determines the post-

pixels consisting of the five pixel elements indicated by the O's. The choice of M is made at the time of impleme ation and is dependent on the am(it of smoothness desired and the type of resize filter chosen: the larger the M the more smoothing applied to the decompressed video. The O's indicate which decompressed intensity pixels are to be used in the local neighborhood.

The memory unit (302) also stores a predetermined plurality of interpolation coefficients (501). The combining circuit (318) typically comprises a resize filter. The combining circuit accesses the memory (302) once. The combining circuit (318) comprises a filter weight determiner (502), a plurality of multipliers (504), and a summer (506). The filter weight determiner (502) determines a plurality of filter coefficients (508). The plurality of multipliers (504) generates a plurality of weighted decompressed pixel intensities (510). The summer provides the post-processed pixel intensity (320).

The post-processing filters (400 and 500) may be implemented using a Digital Signal Processor (DSP) such as the MOTOROLA DSP56002 or an application specific integrated circuit (ASIC) and a Random Access Memory (RAM) unit. The memory unit (302) may be the RAM unit. The signal variance determiner (310), the optimal filter gain (314), and the predictor unit (402) may be implemented using the ASIC. The updater may be implemented using two adders (414 and 418) and a multiplier (416) within the ASIC.

The present invention provides a method and apparatus for removing both mosquito and blocking artifacts from a decompressed video sequence. With such a method and apparatus, it is possible to simultaneously resize the decompressed video while removing these artifacts. This method and apparatus has the following advantages over present post-processing filters: it is independent of signal-to-noise calculations, which are poor indicators of artifact locations; it can perform resizing and post-processing with a single call to memory; a higher perceived viewing quality is obtained.

Although exemplary embodiments are described above, it will be obvious to those skilled in the art that many alterations and modifications may be made without departing from the invention. Accordingly, it is intended that all such alterations and modifications be included within the spirit and scope of the invention as defined in the appended claims.

processed pixel intensity (320) based on the predetermined autocorrelation coefficient (306), the local n ghborhood of pixels (308), and th ptimal filter gain (316). The combining circuit (318) may be implemented as an infinite impulse response (IIR) filter or a finite impulse response (FIR) filter.

FIG. 4, numeral 400, is a block diagram of a post-processing filter which suppresses mosquito and blocking artifacts, using a predictor (402) and an updater (404), in accordance with the present invention. The post-processing filter comprises a memory unit (302), a signal variance determiner (310), an optimal filter gain determiner (314), and a combining circuit (318). The memory unit (302) receives and stores a predetermined noise variance (304), a predetermined autocorrelation coefficient (306), and a decompressed video sequence (216). The decompressed video sequence (216) is stored as a local neighborhood of pixels (308). It should be noted that FIG. 4 reflects the case when M=2 which corresponds to the local neighborhood of pixels consisting of the five pixel elements indicated by the X's and O's. The choice of M is made at the time of implementation and is dependent on the amount of smoothness desired: the larger the M the more smoothing applied to the decompressed video. The X's and O's indicate which pixels are previously post-processed intensity pixels, X's, and which are decompressed intensity pixels, O's, assuming a left to right movement across each row of the image frame. Other methods can also be used for moving through the decompressed image frame stored in memory, such as from right to left across each row, top to bottom or bottom to top across each column.

The general operation of the block diagram shown in FIG. 4 is the same as the block diagram in FIG. 3 and more detail is shown for one implementation of the combining circuit (318). The combining circuit (318) comprises a predictor (402) and an updater (404). The predictor (402) predicts the intensity of the current pixel. The updater (404) determines the post-processed pixel intensity.

The predictor (402) uses a one-dimensional non-causal auto-regressive model of the local neighborhood of pixels (308). An adder (408) adds the first previous post-processed pixel (yb (n-1)) with the first decompressed pixel (x(n+1)) past the pixel under consideration (x(n)). The resultant of this operation is the sum of the current intensity levels (409) at these two pixel locations. The coefficient (407) provided by first order linear prediction coefficient determiner (406) is then used to multiply (410) the sum of intensities (409) . This results in the estimate (412) of the original intensity at pixel location n. The estimate (412) is then compared (414) to the decompressed pixel intensity (x(n)). The output of this comparison (415) is then multiplied (416) by the output (316) of the optimal filter gain determiner (314). The output of this value (417) is summed with the estimate of the intensity value (412), using an adder (418), resulting in the post-processed pixel intensity (320).

FIG. 5, numeral 500, is a diagram of a post-processing filter which suppresses mosquito and blocking artifacts and performs resizing in accordance with the present invention. The post-processing filter comprises a memory unit (302), a signal variance determiner (310), an optimal filter gain determiner (314), and a combining circuit (318). The memory unit (302) receives and stores a predetermined noise variance (304), a predetermined autocorrelation coefficient (306), and a decompressed video sequence (216). The decompressed video sequence (216) is stored as a local neighborhood of pixels (308). As in FIG. 4, FIG. 5 reflects the case when M=2 which corresponds to the local neighborhood of